

Développement : Théorème des extrema liés

Leçons : 151, 159, 214, 215, 217, 219.

Théorème. Soit U ouvert de \mathbb{R}^n , soient $f, g_1, \dots, g_r \in C^1(U, \mathbb{R})$, soit $\Gamma = \{x \in U, g_1(x) = \dots = g_r(x) = 0\}$. Si $f|_{\Gamma}$ possède un extremum relatif en $a \in \Gamma$, et si les formes linéaires $(dg_1(a), \dots, dg_r(a))$ forment une famille libre,

alors il existe $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ tels que $df(a) = \sum_{i=1}^r \lambda_i dg_i(a)$.

Démonstration.

Étape 1 : Cadre de résolution

Comme $(dg_1(a), \dots, dg_r(a))$ famille libre de $(\mathbb{R}^n)^*$ qui est de dimension n , on a $r \leq n$. De plus, si $r = n$, le théorème est évident, les $(dg_i(a))_i$ formant alors une base de $(\mathbb{R}^n)^*$. On suppose donc $r \leq n - 1$, on pose $s = n - r$ et on identifie \mathbb{R}^n à $\mathbb{R}^s \times \mathbb{R}^r$. Les éléments de \mathbb{R}^n seront ainsi notés $(x, y) = (x_1, \dots, x_s, y_1, \dots, y_r)$, $x \in \mathbb{R}^s, y \in \mathbb{R}^r$. On note $a = (\alpha, \beta)$, $\alpha \in \mathbb{R}^s, \beta \in \mathbb{R}^r$.

Étape 2 : Expression locale des éléments de Γ

Comme $(dg_1(a), \dots, dg_r(a))$ est libre, elle est de rang r , et donc la matrice

$$A = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \dots & \frac{\partial g_1}{\partial x_s}(a) & \frac{\partial g_1}{\partial y_1}(a) & \dots & \frac{\partial g_1}{\partial y_r}(a) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1}(a) & \dots & \frac{\partial g_r}{\partial x_s}(a) & \frac{\partial g_r}{\partial y_1}(a) & \dots & \frac{\partial g_r}{\partial y_r}(a) \end{pmatrix} \in \mathcal{M}_{r,n}(\mathbb{R})$$

est aussi de rang r . Comme le rang d'une matrice est la dimension maximale de ses sous-matrices inversibles, on peut extraire de A une sous-matrice carrée inversible de taille r . Quitte à échanger les variables, on peut supposer que $\left(\frac{\partial g_i}{\partial y_j}(a)\right)_{1 \leq i, j \leq r}$ est inversible. Ainsi, si $g = (g_1, \dots, g_r)$, cela revient à dire que $d_y g(a)$ est inversible, on peut donc appliquer le théorème des fonctions implicites : il existe V voisinage de α dans \mathbb{R}^s, W voisinage de β dans \mathbb{R}^r et $\varphi \in C^1(V, W)$ tels que $V \times W \subseteq U$ et $\forall (x, y) \in V \times W, g(x, y) = 0 \Leftrightarrow y = \varphi(x)$. Ainsi, au voisinage de a , $(x, y) \in \Gamma \Leftrightarrow y = \varphi(x)$, en particulier $a = (\alpha, \varphi(\alpha))$.

Étape 3 : Condition d'optimalité

Pour $x \in V$, on pose $h(x) = f(x, \varphi(x))$. h admet un extremum local en α , on a donc $dh(\alpha) = 0$. Or $h = f \circ \psi$ où $\psi = (\text{Id}_{\mathbb{R}^s}, \varphi)$. Par dérivation composée, cela entraîne

$$\forall 1 \leq k \leq s, 0 = \frac{\partial h}{\partial x_k}(\alpha) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\psi(\alpha)) \frac{\partial \psi_j}{\partial x_k}(\alpha) = \sum_{j=1}^s \frac{\partial f}{\partial x_j}(\psi(\alpha)) \frac{\partial \psi_j}{\partial x_k}(\alpha) + \sum_{j=s+1}^n \frac{\partial f}{\partial x_j}(\psi(\alpha)) \frac{\partial \psi_j}{\partial x_k}(\alpha)$$

Or pour $1 \leq j \leq s, \frac{\partial \psi_j}{\partial x_k}(\alpha) = \delta_{j,k}$. D'où, en notant que $\psi(\alpha) = a$ et $x_{s+j} = y_j$,

$$\forall 1 \leq k \leq s, 0 = \frac{\partial h}{\partial x_k}(\alpha) = \frac{\partial f}{\partial x_k}(a) + \sum_{j=1}^r \frac{\partial f}{\partial y_j}(a) \frac{\partial \varphi_j}{\partial x_k}(\alpha)$$

Par ailleurs, en différentiant l'expression $g(x, \varphi(x)) = 0$ en $x = \alpha$, on tire de même

$$\forall 1 \leq i \leq r, \forall 1 \leq k \leq s, 0 = \frac{\partial (g_i \circ \psi)}{\partial x_k}(\alpha) = \frac{\partial g_i}{\partial x_k}(a) + \sum_{j=1}^r \frac{\partial g_i}{\partial y_j}(a) \frac{\partial \varphi_j}{\partial x_k}(\alpha)$$

Alors, si on pose

$$M = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \dots & \frac{\partial f}{\partial x_s}(a) & \frac{\partial f}{\partial y_1}(a) & \dots & \frac{\partial f}{\partial y_r}(a) \\ \frac{\partial g_1}{\partial x_1}(a) & \dots & \frac{\partial g_1}{\partial x_s}(a) & \frac{\partial g_1}{\partial y_1}(a) & \dots & \frac{\partial g_1}{\partial y_r}(a) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1}(a) & \dots & \frac{\partial g_r}{\partial x_s}(a) & \frac{\partial g_r}{\partial y_1}(a) & \dots & \frac{\partial g_r}{\partial y_r}(a) \end{pmatrix} \in \mathcal{M}_{r+1,n}(\mathbb{R})$$

les relations précédentes montrent que les s premières colonnes de M s'expriment comme combinaison linéaire des r dernières. Donc $\text{rg}(M) \leq r$. De plus, comme le rang des lignes est égal au rang des colonnes (car $\text{rg}({}^tM) = \text{rg}(M)$), les $r + 1$ vecteurs lignes de M forment une famille liée : il existe $\mu_0, \dots, \mu_r \in \mathbb{R}$ non tous nuls tels que $\mu_0 df(a) + \sum_{i=1}^r \mu_i dg_i(a) = 0$. Comme $(dg_1(a), \dots, dg_r(a))$ libre, $\mu_0 \neq 0$, et en posant pour tout i , $\lambda_i = -\mu_i/\mu_0$, on a $df(a) = \sum_{i=1}^r \lambda_i dg_i(a)$. □

Compléments

Une caractérisation de $SO_n(\mathbb{R})$

Théorème. $SO_n(\mathbb{R})$ est l'ensemble des éléments de $SL_n(\mathbb{R})$ de norme minimale.

Démonstration.

On pose $g(M) = \det(M) - 1$, $N(M) = \text{tr}({}^tMM) = \|M\|^2$. On cherche à minimiser N sous la contrainte $g(M) = 0$. On a $N(A + H) = N(A) + 2\text{tr}({}^tAH) + \text{tr}({}^tHH)$ donc N différentiable et $dN(A)H = 2\text{tr}({}^tAH)$. De plus, si C est la comatrice de A , on a $dg(A)H = \text{tr}({}^tCH)$. Enfin, comme $SL_n(\mathbb{R}) = g^{-1}(\{0\})$ est fermé, N atteint son minimum sur $SL_n(\mathbb{R})$ car la distance δ de 0 à un fermé F est toujours atteinte (sur le compact $F \cap \overline{B}(0, \delta + 1)$).

D'après le théorème des extrema liés, si N atteint son minimum en A , alors $\forall H, 2\text{tr}({}^tAH) = \lambda \text{tr}({}^tCH)$. En évaluant cette relation en tous les $H = E_{i,j}$, on tire $2A = \lambda C$. Comme A inversible et $\det(A) = 1$, on a $C = {}^tA^{-1}$ d'où $2 {}^tAA = \lambda I_n$. En composant par le déterminant, il vient $\lambda^n = 2^n$ donc comme ${}^tAA \in S_n^+(\mathbb{R})$, $\lambda > 0$ donc $\lambda = 2$. Donc ${}^tAA = I_n$ et $A \in SO_n(\mathbb{R})$. Le minimum de N vaut donc $N(I_n) = n$. Réciproquement, si $A \in SO_n(\mathbb{R})$, $N(A) = n$. □

Références

X. Gourdon, LES MATHS EN TÊTE - ANALYSE, Ellipses, p. 317, 321, 327.